Online Services • BY KEVIN M. SAVETZ

Here Come The Knowbots!

Although it sounds like a character from a William Gibson novel, the concept of the "knowbot," an automated program that searches for information for you, is slowly becoming a reality on the Internet. Knowbots stand to change the face of research and information retrieval: Rather than spending hours searching for an electronic needle in a virtual haystack, why not let the information come to you?

I first wrote about a kind of knowbot last year in a *MicroTimes* article about the Stanford Netnews Filtering Service. It's a program that periodically searches the entire Usenet for the information you request and emails any "hits" back to you. (For information about it, send mail to *netnews@db.stanford.edu* with the word "help" in the message body.)

I'm happy to report that the concept and power of knowbots are expanding on the Net. The actual implementations of today's knowbots are decidedly less sexy than the futuristic-sounding self-replicating, auto-destructing knowbot defined above. And they still don't resemble the fancy animated "agents" that Apple Computer predicted in a futuristic promotional video a few years ago. Still, knowbots are very handy as automated news clipping services. Here we'll examine what you can do with two more knowbot utilities—another one from Stanford and one by the folks at *San Jose Mercury News*.

knowbot: *an independent, self-acting computer program that seeks for information on behalf of a user, possibly replicating itself on other hosts on the network. As the knowbot performs its task, it sends reports back to the user, and self-destructs when it completes its task.*

—Internet in Plain English *by Bryan Pfaffenberger*

Mercury Center NewsHound

The Mercury Center NewsHound is an easy-to-use news clipping service that offers the first low-cost news search agent aimed at consumers. NewsHound automatically searches the stories and classified ads in the *San Jose Mercury News* as well as hundreds of stories not published in the paper. Articles and ads matching your interests are sent directly to you via email.

Here's how it works: Once you sign up, you send a search request via email to the NewsHound program. Every hour or so, NewsHound checks to see if any of its recent news and ads match your request. It automatically sends the ones that do to you via electronic mail. I use the service to send me articles about online services, my favorite musicians, and ads for used computers. One might also use it to scope out editorials about Newt Gingrich, news about your home town, or ads for used Pintos.

Each request that you submit includes lists of "required" terms, "possible" ones, and "excluded" terms—phrases that you just don't want to hear about. The NewsHound uses a type of fuzzy logic to find the most relevant articles and advertisements. Essentially this involves counting the number of "possible" and "required" terms found in an article and then assigning the article a "selectivity" score. The score of an article ranges from 1 to 100. The higher the number, the more relevant the article.

NewsHound searches a variety of newsy databases in pursuit of your information needs, including the *San Jose Mercury News*, the *Chicago Tribune*, *Detroit Free Press*, *Miami Herald*, *Philadelphia Inquirer*, *The New York Times* News Service, The Associated Press, and PR Newswire as well as published and unpublished articles from sixty other sources, for a total of about two thousand articles each day.

NewsHound isn't free, but it's inexpensive enough to be affordable for just about everyone. The special rate for charter subscribers is $4.95 per month for up to five search profiles. That is a flat rate that applies regardless of the number of stories or ads delivered to your emailbox during the month. (The rate for noncharter subscribers will be $9.95 each month.)

To register to use NewsHound, call 1-800-818-NEWS or (408)297-8495. For

continued on page 196

continued from page 120

more information, send electronic mail to *newshound-support@sjmercury.com* .

Computer Science Technical Report Service

As part of the DARPA Electronic Library Project, the Database Group at Stanford provides a free service to disseminate information about computer science technical reports. If you're a computer scientist, listen up. And if you're not, listen anyway; this will still give you an idea of the power of knowbots.

A user interested in receiving bibliographical records of technical reports available from the Math/Computer Science Library at Stanford submits their "interest profile" to the service. Periodically they will receive new information relevant to their interests via electronic mail.

Queries formed as plain English text (for example "high-speed fiber optics communication") are submitted via email. Abstracts of technical reports are returned based on their relevancy to your query. As new and relevant reports are added to the database, the knowbot will inform you. Some of them are available digitally—you can receive the full text with a simple emailed request. For others, you'll have to trek to your local library (or Stanford's) for the actual report. For more information about using the service, send email with the word "help" in the message body to *elib@db.stanford.edu* .

What's Next?

Admittedly, computer science technical reports are pretty dry, and only of interest to a small group, but imagine what might be done with other databases running on knowbot servers. How about a recipe database, updated weekly, that automagically emails you vegetarian recipes featuring red potatoes or asparagus (but not beets)? Or an archive of *MicroTimes* articles that sends you everything mentioning Macintosh, except for stuff by that pesky John Dvorak?

Stay tuned. I predict those things (and more) are coming to the Internet, sooner or later.■

Copyright © 1995 by Kevin Savetz. All rights reserved.

Example Of NewsHound Service

To: NewsHound@sjmercury.com
Subject: request
create: music
search: articles
possible: laurie anderson,negativland,they might be giants
selectivity: 20

From: NewsHound@sjmercury.com (NewsHound)
To: savetz@northcoast.com
Subject: Your NewsHound request
You now have 3 NewsHound profiles. Their attributes are:
Title: INTERNET
Search: articles
Possible: internet,on-line service,online service,worldwideweb,mosaic,compuserve,america online,eworld,prodigy
Excluded: information superhighway
Selectivity: 30

Title: MUSIC
Search: articles
Possible: laurie anderson,negativland,they might be giants
Selectivity: 20

Title: SUN
Search: ads
Possible: sun,sparc,sparcstation,unix,macintosh,mac
Required: for sale
Selectivity: 50 —

From: NewsHound@sjmercury.com (NewsHound)
To: savetz@northcoast.com
Subject: [30] STUDENT-RUN RADIO STATION A FIRST ON THE INTERNET
Selected by your NewsHound profile entitled "INTERNET". The selectivity score was 30 out of 100.
Student-Run Radio Station a First on The Internet
By JULIANNE BASINGER, Associated Press Writer
CHAPEL HILL, N.C. (AP) — The campus radio station at the University of North Carolina is serving up its musical fare to a wider audience these days: Listeners abroad can jam along — if they have the right computer.

WXYC-FM recently jumped onto the Internet, becoming the first radio station to offer 24-hour programming on the air and the global computer network simultaneously.

People who want to listen can do so by gaining access to the university's site on the World Wide Web, a part of the Internet that has become popular because of the development of sound, video and graphical features. ...

Example Of An Electronic Library Project Search

From: Electronic Library <elib@DB.Stanford.EDU>
To: savetz@northcoast.com
Subject: Elib Response
Search: knowbots information filtering
Threshold: 50
Score: 66
ID: 124547
Location: Math/CS Library
Author: Goldberg, D.
Author: Nichols, D.
Author: Oki, B.
Author: Terry, D.
Title: Using collaborative filtering to weave an information tapestry.
Institute: Xerox Corporation. Palo Alto Research Center.
Report: CSL-92-10.
Date: 1992.
Keyword: Continuous query.
Keyword: Information filtering.
Keyword: Active database.
Keyword: Query rewrite.
Keyword: Bounding monotonic query.
Keyword: Incremental query.
Keyword: Query language.
Score: 59
ID: 025494
Location: Math/CS Library
Author: Yan, Tak W.
Author: Garcia-Molina, Hector.
Title: Index structures for information filtering under the vector space model.
Institute: Stanford University. Department of Computer Science.
Report: STAN-CS-TR-93-1494.
Date: 1993.
Score: 55
ID: 124542
Location: Math/CS Library
Author: Terry, D.
Author: Goldberg, D.
Author: Nichols, D.
Author: Oki, B.
Title: Continuous queries over append-only databases.
Institute: Xerox Corporation. Palo Alto Research Center.
Report: CSL-92-5.
Date: 1992.
Keyword: Information filtering.
Keyword: Active database.
Keyword: Query rewrite.
Keyword: Bounding monotonic query.
Keyword: Incremental query.